Music Genre Classification

Name:	**Nitish Kumar**
Registration No./Roll No.:	17171
Institute/University Name:	IISER Bhopal
Program/Stream:	BS-MS
Problem Release date:	February 02, 2022
Date of Submission:	April 24, 2022

1 Introduction

A music genre is a conventional category that identifies some pieces of music as belonging to a shared tradition or set of conventions. The various genres of music are : country, disco, hiphop, pop, classical, blues, rock, metal, jazz, or reggae.

The music label distribution is almost in the training as it can be seen from 1. It can be seen here that the all the features in the data set are not normally distributed and some of the variance features are highly left or right skewed. We will try to perform feature transformations so that these features have normal distribution.

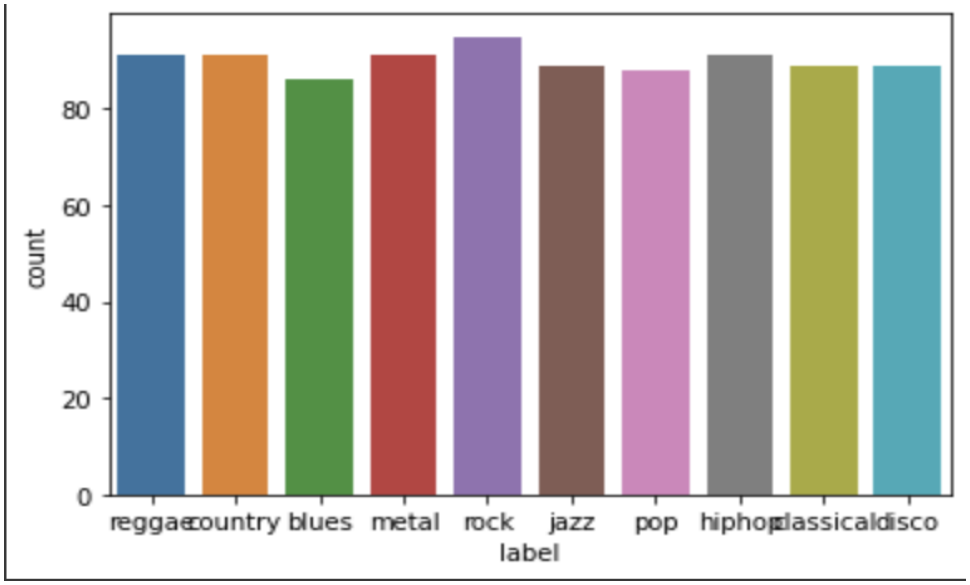

Figure 1: Count of various labels in the training dataset

The training set comprises of 900 data points and there 58 features corresponding to each data points and there are 10 different genres. The various features of the training dataset are not on the same scale. So, we will proceed by normalization and then apply various machine learning algorithms to get our predictions.

2 Methods

We initially started by working with raw data (as it is given to us) but the results obtained by Random Forest Classifier (RFC) and SVC were very poor. So, we started by normalizing the features in the train and test data sets. Then, as we have already noted that the variance columns were left/right

Table 1: Performance Of Different Classifiers Using All Terms on Validation Set

Classifier	Accuracy
Random Forest Classifier	0.778
SVC	0.76
KNN	0.687
Multinomial Logistic Regression	0.726
LGBM	0.8011
XGB	0.7911

Table 2: Performance Of Different Classifiers after Feature Selection

Classifier	Accuracy
Random Forest Classifier	0.78
SVC	0.743
KNN	0.713
Multinomial Logistic Regression	0.702
LGBM	0.782
XGB	0.775

skewed, we proceed to make them normalized. To do that, we made use of box-cox transformation and by doing fit_transform on train and transform on test data set, by doing this we were able to normalize them. We work using these transformed data sets for further study.

The following machine learning models were used: Random Forest Classifier (RFC), SVC, KNN, Multinomial Logistic regression, Light Gradient Boosting Machine (LGBM)[1] and XGBoost (XGB) [2].

In order to achieve hypertuning, we made use of *Optuna*. Optuna is an automatic hyperparameter optimization software framework, particularly designed for machine learning. One of the advantages of Optuna is that it allows us to define objects that define search spaces dynamically.[3]

We then explore feature selection technique using *mutual_info_classif* and we make use of *SelectKBest* to select top 30 features having the best mutual information. After that we try Principal Component Analysis (PCA) by setting the hyperparameter *n_components* to 0.95 since we want the explained variance to be around 95%.

After getting the predictions, I used the ensemble technique known as *bagging method*. For classification problem, bagging method can be interpreted as the majority of the decisions of several models.

For the bagging method, I stacked 13 models in total: the top 5 models which had the best accuracy from Table 1, top 5 models from Table 2 and top 3 models from Table 3, along their rows. And then I took mode (highest frequency) along the rows in order to achieve the majority voting from different models in order to get my final predictions.

The Github link for the project can be accessed here.

3 Evaluation Criteria

In this project, we worked with Accuracy as our evaluation criteria. Accuracy denotes the proportion of the total number of predictions that were correct. It is one of the simplest form of evaluation metrics.

$$\text{Accuracy} = \frac{\text{number of correctly classifier data points}}{\text{number of total data points}}$$

Table 3: Performance Of Different Classifiers after PCA

Classifier	Accuracy
Random Forest Classifier	0.718
SVC	0.7566
KNN	0.704
Multinomial Logistic Regression	0.7077
LGBM	0.701
XGB	0.7067

4 Analysis of Results

We have created three tables, namely Table 1, Table 2 and Table 3. Table 1 denotes the accuracy of various machine learning models after normalization and transformation of data sets, Table 2 denotes the accuracy after performing feature selection and Tbale 3 denotes the accuracy of various models after performing PCA.

It can be seen that after performing Feature selection, the accuracy of some models were increased and for some the accuracy was decreased. Also note that, PCA doesn't yield any sort of significance improvement. This might be due to the fact that almost all features are important and by doing PCA we might have lost some crucial information which might be a important factor for correct prediction.

5 Discussions and Conclusion

If we try to understand the different genres, one can note that:

- Reggae incorporates some of the musical elements blues and jazz.
- Hip hop music was influenced by disco music.

Hence, there are might be some key information due to which some of our points might be misclassified. Also doing PCA took away some crucial information due to which the performance of the models got decreased and hence it was not a suitable idea where as Feature selection was able to preserve those features and eventually led to almost same or increased performance.

We can improve our results by taking care of the long tails that appear in the variance features by doing something different than box-cox transform, for e.g. negative binomial transformations. Also, since the data set was small and each class label had almost 90 data points, it is clear that the models were not able to learn to the best of their ability. In case, we had a larger data set, it seems the models would have performed better. Also, the feature set only contained mean and variances of different music characteristics, which might have led to wrong predictions due to some close relations between different genres of music.

References

[1] *Light Gradient Boosting Machine*

[2] *XGBoost*

[3] *Optuna*